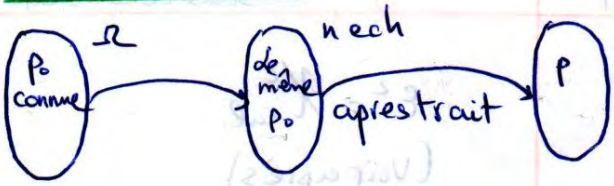
Free database on www.la-faculte.net published for NON-lucrative use

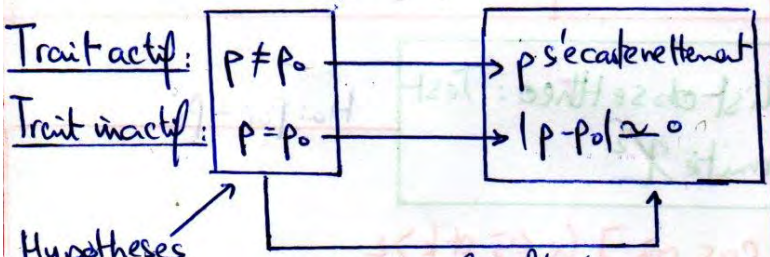
② Tests d'Hypotheses

A/ Generalités sur les Tests

Introduction:



On veut tester l'efficacité du traitement
⇒ On compare p et p_0



Principe General des Tests

→ Definir H_0 et H_1 : (H_0 ⇒ l'egalité)

→ Calcul: Construire une stat T d'une loi connue + calculs.

→ Conclusion: Rejetter ou ne pas rejeter H_0 avec un risque α.

Qlq parametres à connaitre:

→ erreur de type I : $\alpha = P(\text{rejeter } H_0 / H_0 \text{ vraie})$

→ erreur de type II : $\beta = P(\text{rejeter } H_1 / H_1 \text{ vraie})$

$\alpha + \beta \neq 1$ et α est plus grave

→ Puissance : $\Pi = 1 - \beta = P(\text{conclure } H_1 / H_1 \text{ vraie})$

→ La p-value: * erreur max → ne pas rejeter H_0
* erreur min → rejeter H_0

$$\text{pvalue} = \begin{cases} 2F(t_0) & \text{si } F(t_0) \leq 0,5 \\ 2(1 - F(t_0)) & \text{si } F(t_0) \geq 0,5 \end{cases}$$

B/ Tests de Conformité

* (ech - pop)
⇓
le parametre à etudier est connu
* H_0 : ech representatif de la pop.

Comparaison d'1 moy obs $\bar{x}$ et 1 moy theo connu m

Intervalle de Parv

- $I P_\alpha(m) = [m \pm h]$; h selon les ≠ cas.
- Si $\bar{x} \notin I P_\alpha(m)$ ⇒ H_0 rejetée au risque α.

l'ecart reduit ε:

- $\varepsilon = \frac{\sqrt{n}}{\sigma} |\bar{X} - m| \rightsquigarrow N(0,1)$ si σ connue
- $\varepsilon = \frac{\sqrt{n-1}}{s} |\bar{x} - m|$ si σ inconnue.

$n \leq 30$: St_{n-1} On suppose la normalité

$n > 30$: $N(0,1)$

- $\varepsilon > q_{1-\alpha/2}$ ou $t_{n-1}(\alpha)$ ⇒ H_0 rejetée au risque α.

NAIDJI ADEL 2015/2016

Contact us on: facadm16@gmail.com 2015/2016

Free database on www.la-faculte.net published for NON-lucrative use

Comparaison d'1 prop obs p et 1 prop theo p_0 connue

Conditions de validité : $n > 30$, $np_0 \geq 5$ et $nq_0 \geq 5$

IP	E	Conformité du χ^2
Les 3 condit°	Les 3 condit°	Sans condition
$IP_\alpha(p_0) = [p_0 \pm h]$ $h = q_{1-\alpha/2}\sqrt{\frac{p_0 q_0}{n}}$ $p \notin IP \Rightarrow H_0$ rejeté au risque α	$\varepsilon = \frac{\lvert p - p_0 \rvert}{\sqrt{\frac{p_0 q_0}{n}}} \sim N(0,1)$ $\varepsilon > q_{1-\alpha/2} \Rightarrow H_0$ rejeté (α)	$\varepsilon^2 = \chi^2_{cal}$ (Voir après)

Comparaison de 2 dist obs et theo : Test de Conformité χ^2

$H_0 : p_{oi} = p_i$

Cas general : $C_i \geq 5$:

Classe	$O_i = np_i$	p_{oi}	$C_i = np_{oi}$	$\frac{\lvert O_i - C_i \rvert^2}{C_i}$
Il y a k classes	effectifs	proport°	effectifs	
Σ	n	1	n	χ^2_{cal}
	obs	theorique.		

$$\chi^2_{cal} = \sum_{i=1}^{k} \frac{(O_i - C_i)^2}{C_i} \quad \forall i = \overline{1,k} \quad C_i \geq 5.$$

si $\chi^2_{cal} > \chi^2_{\upsilon}(\alpha)$: H_0 rejeté (α)

$\upsilon = k - 1$

Cas où $\exists C_i \leq 5$ et $k > 2$

On regroupe [la classe avec C_i plus petit] avec [la plus petite classe adjacente] jusqu'à dépasser 5.

et $\upsilon = k' - 1$

Cas où $\exists C_i < 5$ et $k = 2$:

Correction de Yates

$$\chi^2_{cal} = \sum_{1}^{2} \frac{\left(\lvert O_i - C_i \rvert - \frac{1}{2}\right)^2}{C_i}$$

et $\upsilon = 1$.

Cas particulier $k = 2$:

$\chi^2_{cal} = \varepsilon^2$

Tests d'Adjustement

* Comparer les distribut°

\+ Peut-on adjuster... ?

=

Test de conformité du χ^2

- p_{oi} est donnée par une loi usuelle : $\upsilon = k - 1$
- p_{oi} est donnée par une loi usuelle : $\upsilon = k - 1 - r$.

Contact us on: facadm16@gmail.com 2015/2016

Free database on www.la-faculte.net published for NON-lucrative use

Lois	Calcul des p_{oi}	des classe	r = nbre de parametres inconnus = estimés.
Binomiale $B(n,p)$	$p_{oi} = C_n^i p^i (1-p)^{n-i}$	$i \in \{0, 1, \dots, n\}$	$r = \{0,1\}$ et $\hat{p} = \frac{\bar{x}}{n}$
Poisson $P(\lambda)$	$p_{oi} = e^{-\lambda}\frac{\lambda^i}{i!}$ $\begin{cases} P(X=i+1) = \frac{\lambda}{i+1} P(X=i) \\ P(X=0) = e^{-\lambda} \end{cases}$	$i = \{0, 1, \dots, \infty\}$ $i > k$ $O_k = 0$ $C_k = n p_{ok}$	$r = \{0, 1\}$ et $\hat{\lambda} = \bar{x}$
Normale $N(m, \sigma^2)$	La table de $N(0,1)$	On ajoute 2 classes $]-\infty, a_0[$ et $]a_k, +\infty[$ $O_i = 0$ | $C_i = n p_{oi}$ | $O_i = 0$ | $C_i = n p_{oi}$	$r = \{0,1,2\}$ et $\hat{m} = \bar{x}$, $\hat{\sigma}^2 = \frac{n}{n-1} s^2$

$\bar{x} = \frac{1}{n}\sum n_i x_i$

$s^2 = \frac{1}{n}\sum n_i x_i^2 - \bar{x}^2$

S'il s'agit d'une normale x_i, c'est le centre de l'intervalle

C/ Tests d'homogeneité

$(ech_1 - ech_2)$ H_0 : les 2 ech proviennent de 2 pop identiques / il y a independance.

Comparaison de 2 moy

$H_0 : m_1 \simeq m_2$ / $\bar{x}_1 \simeq \bar{x}_2$

Recouvrement par IC

$I_1 = IC_\alpha(m_1)$ et $I_2 = IC_\alpha(m_2)$

- Si $I_1 \cap I_2 = \emptyset$: H_0 rejetée (α)
- Si $I_1 \cap I_2 \neq \emptyset$
  - $\bar{x}_1$ et $\bar{x}_2 \in I_1 \cap I_2$ → H_0 non rejetée
  - $\bar{x}_1$ ou $\bar{x}_2 \notin I_1 \cap I_2$ → On peut pas juger

L'ecart reduit ε

n_1 et $n_2 > 30$:

$$\varepsilon = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \rightsquigarrow N(0,1)$$

$\varepsilon > q_{1-\alpha/2} \Rightarrow H_0$ rejetée.

n_1 ou $n_2 \leq 30$:

$$\varepsilon = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} \rightsquigarrow St_{n_1+n_2-2}$$

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$$

$\varepsilon > t_{n_1+n_2-2} \Rightarrow H_0$ rejetée (α)

- Si $n_1 + n_2 - 2 > 30 \Rightarrow$ On approxime par N

NAIDJI ADEL 2015/2016

Comparaison de 2 prop

IC $n_1 > 30$, $n_2 > 30$	$n_1 p_1, n_1 q_1, n_2 p_2, n_2 q_2 \geq 5$ — ε	χ^2
kif kif	$\varepsilon = \frac{\lvert p_1 - p_2 \rvert}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$; $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ $\varepsilon \rightsquigarrow N(0,1)$ $\varepsilon > q_{1-\alpha/2} \Rightarrow H_0$ rejetée	$\varepsilon^2 = \chi^2_{cal}$ (voir après)

Contact us on: facadm16@gmail.com 2015/2016

Free database on www.la-faculte.net published for NON-lucrative use

ça concerne la moy bcp plus. **Taille minimale** pour H_0 rejeté.

IC $I_1 \cap I_2 = \Phi$	ε
On suppose $n_1 = n_2 = n$. $\overline{x_1}+h$ $\overline{x_2}-h$ $\overline{x_1}+h_1 < \overline{x_2}-h_2$ / h est une fonct° de n $n > \dots \left(q_{1-\alpha/2}\dfrac{S_2+S_1}{\overline{x_2}-\overline{x_1}}\right)^2+1$	$\varepsilon > q_{1-\alpha/2}$; ε est une fonct° de n. $n > \dots \left(q_{1-\alpha/2}\dfrac{\sqrt{S_2^2+S_1^2}}{\overline{x_2}-\overline{x_1}}\right)^2$. Meilleure Methode $\Rightarrow$ celle qui nous donne n↘

D/ Test d'independance du χ^2

(de lien entre 2 caracters qualitatifs)

H_0 : Il y a Independance.

Cas general $C_{ij} \geq 5$:

ex : Influence du Tabac T sur le Cancer C

T \ C	C (1)	$\overline{C}$ (2)	Σ
T (1)	$O_{11}=40$ / C_{11}	$O_{12}=30$ / C_{12}	70
$\overline{T}$ (2)	$O_{21}=20$ / C_{21}	$O_{22}=10$ / C_{22}	30
Σ	60	40	Σ 100

- O_{ij} : effectifs observes
- C_{ij} : effectifs theoriques $= \dfrac{\Sigma \times \Sigma}{\Sigma}$

$$\begin{cases} C_{11} = \dfrac{60\times 70}{100} \\ C_{21} = \dfrac{60\times 30}{100} \dots \end{cases}$$

ici, on a 4 classes, chacune a un O_{ij} et C_{ij}

x_{ij}	O_{ij}	C_{ij} (calculés	$(O_{ij}-C_{ij})^2/C_{ij}$
Σ	n	n	χ^2_{cal}

$$\chi^2_{cal} = \sum_i \sum_j \frac{(O_{ij}-C_{ij})^2}{C_{ij}}$$

si $\chi^2_{cal} > \chi^2_{\nu}(\alpha) \Rightarrow H_0$ rejetée (α)

$$\boxed{\nu = (k-1)(l-1)}$$

- **Cas où $\exists\, C_{ij} < 5$ et $k > 2$:** (On regroupe les caracteres influencés)

regroupement de classes $\nu = (k'-1)(l'-1)$

- **Cas où $\exists\, C_{ij} < 5$ et $(k,l) = (2,2)$:**

$$\chi^2_{yates} = \sum_1^2 \sum_1^2 \frac{\left(|O_{ij}-C_{ij}| - \frac{1}{2}\right)^2}{C_{ij}}$$

$\nu = 1$

- **Cas particulier $(k,l) = (2,2)$:**

2 methode

$$\chi^2_{cal} = \varepsilon^2$$

Test du χ^2 — Comparaison de p_1 et p_2

$h_1 = \sqrt{\dfrac{p_1 q_1}{|F|}}$ $\quad p_1 = P(C/F) = \dfrac{|C \cap F|}{|F|} = \dfrac{40}{70}$

$h_2 = \sqrt{\dfrac{p_2 q_2}{|\overline{F}|}}$ $\quad p_2 = P(C/\overline{F}) = \dfrac{|C \cap \overline{F}|}{|\overline{F}|} = \dfrac{20}{30}$.

NAIDJI ADEL 2015/2016

Contact us on: facadm16@gmail.com 2015/2016

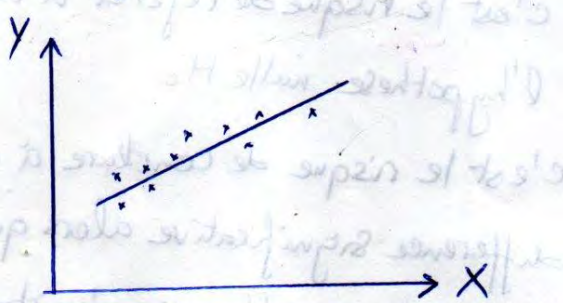
Free database on www.la-faculte.net published for NON-lucrative use

E/ Tests de Corrélation

Introduction:

- Le But est d'étudier l'existence d'un lien entre 2 caractères quantitatifs X et Y
- si ce lien existe, la relation entre X et Y sera considérée linéaire

$$Y = aX + b.$$



nuage de pts (Diagramme de Dispersion)

- Pour trouver (a, b), on doit faire des calculs

Paramètres à Calculer:

La Covariance: $S_{xy} \in \mathbb{R}$

$$S_{xy} = \frac{1}{n}\sum x_i y_i - \bar{x}\bar{y}$$

$$S_{xy} = \frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y})$$

Coefficient de Corrélation:

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$$

$$|r_{xy}| \leq 1$$

- $|r_{xy}| = 1$: Dans la pratique $r_{xy}^2 \geq \frac{3}{4}$ ⇒ X et Y non indép. { $r > 0 / a > 0$ → Y et X croient en sens ; $r < 0 / a < 0$ ⇒ inversement.
- $r_{xy} = 0$: X et Y indép
- Sinon : On peut pas juger ⇓ Tests. (F)

Tests de Corrélation:

H_0: Il y a indépendance

Conformité (1 éch)	Homogeneité (2 éch)
$H_0: r_{xy} \simeq 0$	$H_0: r_1 \simeq r_2$
$\varepsilon = \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \lvert r \rvert \hookrightarrow St_{n-2}$	$\varepsilon = \frac{\lvert r_1 - r_2 \rvert}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \hookrightarrow N(0,1)$
$\varepsilon \geq t_{n-2} \Rightarrow H_0$ rejetée (α)	$\varepsilon \geq q_{1-\alpha/2} \Rightarrow H_0$ rejetée (α)

↳ si $n-2 \geq 30$: $t_{n-2} \simeq q_{1-\alpha/2}$ mais la loi exacte reste me St_{n-2}

Droites de Régression:

Grâce aux paramètres déjà calculés, on peut désigner a, b, α, β.

Dy(x): $\hat{y} = \hat{a}x + \hat{b}$

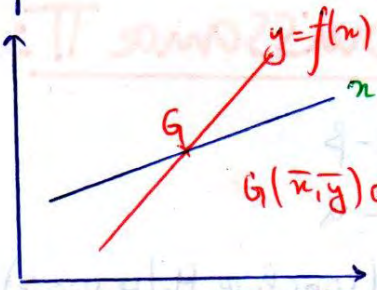
$$\begin{cases} \hat{a} = \frac{S_{xy}}{S_x^2} \\ \hat{b} = \bar{y} - \hat{a}\bar{x} \end{cases}$$

utilité: prédire y_0 d'une valeur obs x_0

Dx(y): $\hat{x} = \hat{\alpha}y + \hat{\beta}$

$$\begin{cases} \hat{\alpha} = \frac{S_{xy}}{S_y^2} \\ \hat{\beta} = \bar{x} - \hat{\alpha}\bar{y} \end{cases}$$

utilité : prédire x_0 d'une valeur obs y_0.



NAIAJI ADEL 2015/2016

Contact us on: facadm16@gmail.com 2015/2016